陇中黄土丘陵沟壑区人工草地土壤水蚀预测模型①

王　钧[1,2]，　李　广[1]，　聂志刚[2]，　刘　强[2]
（1　甘肃农业大学林学院，甘肃　兰州　730000；　2　甘肃农业大学信息科学技术学院，甘肃　兰州　730000）

摘　要：针对陇中黄土丘陵沟壑区土壤水蚀过程复杂且难以有效预测的问题，以定西市安家沟水土保持试验站2005—2016年1～12月人工草地径流场试验数据为主要来源，将流域月降雨量、月侵蚀性降雨量、月径流量、月降雨强度、径流场面积、径流场坡度、土壤砂粒含量、土壤粘粒含量8个因子作为输入因子，月土壤水蚀量作为输出，运用偏最小二乘法（Partial Least-Squares Regression，PLSR）和长短期记忆（Long Short-Term Memory，LSTM）循环神经网络建立人工草地土壤水蚀预测模型，并利用BP（Back Propagation）、RNN（Recurrent Neural Network）、LSTM常见神经网络模型，对模型的有效性进行评估。结果表明：PLSR将模型8个输入因子减少为4个，从而有效解决LSTM神经网络模型对样本数量要求过高的问题；PLSR和LSTM神经网络模型的结合可以有效提高模型对人工草地土壤水蚀过程的预测精度和收敛速度，预测结果的平均相对误差小于4%，相关系数高于其他3种神经网络模型，而迭代次数、均方根误差和平均绝对误差均低于其他3种模型；研究发现坡度对人工草地土壤水蚀过程影响较为明显，降雨量小于25 mm时，人工草地土壤水蚀量不会随坡度增加而明显增长，但当降雨量超过25 mm时，人工草地土壤水蚀量会随坡度明显增加。PLSR-LSTM神经网络土壤水蚀预测模型可以准确预测陇中黄土丘陵沟壑区人工草地土壤水蚀量，为该地区水土流失的准确预报提供新的思路和方法。

关 键 词：　黄土丘陵沟壑区；人工草地；长短期记忆循环神经网络；土壤水蚀预测
文章编号：

陇中黄土丘陵沟壑区是我国土壤水蚀最严重的地区之一，土壤水蚀造成该地区土地水土流失严重、沟壑面积增多、农田遭受破坏，严重阻碍了当地经济的可持续发展[1]。研究表明降雨是导致坡面土壤水蚀的主导因素，是导致水土流失最重要的外部驱动，而坡度、坡形、以及植被覆盖对水蚀的影响也极为显著，整个土壤水蚀过程相对复杂，难以做出有效的预测[2－5]。目前，国内外学者对土壤的水蚀机理和规律进行了大量的研究工作，并在此基础上建立了适用于不同地区的土壤水蚀模型，主要包括经验统计模型和物理模型，其中经验统计模型多以多元回归模型为主，模型计算简便，但预测精度不够理想；而物理模型的建立多以土壤水蚀机理为基础，模型虽能较好的反映土壤水蚀过程和机制，但计算较为复杂[6－9]。

近几年，随着深度学习技术的不断成熟，一些国内外研究学者将基于深度学习的各种模型应用于时间序列数据的预测中[10]。深度学习模型是一种由很多非线性隐藏层组成的深度神经网络模型，能够根据输入数据逐层提取数据的有效特征，从而挖掘出数据间的潜在规律[11]。循环神经网络模型（Recurrent Neural Network，RNN）将时间序列的概念引入到神经网络模型的设计中，使其更适合应用于时间序列数据的处理[12]。在深度学习技术不断发展的过程中出现了许多RNN模型的变体，其中长短期记忆（Long Short-Term Memory，LSTM）神经网络模型通过引入“门”解决了RNN长期记忆能力不足的问题，使LSTM神经网络模型真正实现了长距离时间序列数据的有效利用[13]。但模型自变量间的多重相关性以及对样本点个数要求过高等问题，限制

①　**收稿日期：**2019－04－10；　**修订日期：**2019－09－29
基金项目：甘肃农业大学学科建设基金（GAU-XKJS-2018-254，GAU-XKJS-2018-258）；国家自然科学基金项目（31660348）；甘肃省重点研发计划（18YF1NA070）资助
作者简介：王钧（1982－），男，甘肃白银人，博士研究生，讲师，从事水土保持与荒漠化防治研究. E-mail：julianwong82@163.com
通讯作者：李广，男，博士，教授，从事水土保持与荒漠化防治研究. E-mail：lig@gsau.edu.cn

了该模型的实际应用。偏最小二乘法(Partial Least-Squares Regression, PLSR)[14]既充分吸收主成分分析思想[15],同时可以提高主成分之间的相关性[16]。利用该算法可以实现高维数据空间的数据降维,特别适合在数据间存在严重多重相关性以及样本点个数相对较少的模型建模中应用。

本文旨在构建基于 PLSR 优化后的 LSTM(PLSR-LSTM)神经网络人工草地土壤水蚀预测模型,以实现陇中黄土丘陵沟壑区人工草地土壤水蚀的准确预报,为该地区水土流失的准确预测提供有效的技术支撑。

1　研究区概况与方法

1.1　研究区概况

本试验将径流场设置在甘肃省定西市安家沟流域,该地区属于典型黄土高原丘陵沟壑地貌,平均海拔高度 2 250 m,多年平均降雨量为 385 mm,年均蒸发量为 1 531 mm,年平均气温为 6.5 ℃。该地区土壤类型以灰钙土为主,土壤母质为黄土,质地为粉砂壤和粉壤,土质疏松,极易发生土壤水蚀。植被属于森林草原干草原区,人工草地以红豆草和紫花苜蓿为主。

1.2　试验设计

试验所用径流场被设置在安家沟流域阴坡中部,面积为 10 m × 5 m,坡度分别为 5°、10°、15°、20°,每个坡度径流场重复 2 次。径流场 0 ~ 10 cm 土层土壤的砂粒含量和粘粒含量分别为 24.3% 和 9.7%。人工草地覆盖植被为红豆草,覆盖度达到 70% ~85%,平均为 79%。

1.3　试验测定项目和分析方法

(1) 降雨量测定

利用流域内自动气象站和虹吸式雨量计,来记录每次降雨的雨量、降雨时间和平均雨强,以日降雨量≥12 mm 的标准来划分水蚀性降雨量[17]。

(2) 径流量测定

各个径流场通过集流池对径流产沙量进行观测。地表产生径流后,用水尺读取集流池内的水深,再乘以集流池的面积,计算出径流总量。

(3) 土壤水蚀量测定

将集流池中泥水搅拌均匀后,用体积为 500 mL 的取样瓶在集流池的不同区域中取水样 3 次,对取得水样进行过滤后烘干称重,测得水样的泥沙含量,从而得到各径流场的产沙量,产沙量与面积的比值即为土壤水蚀量。

1.4　偏最小二乘法

偏最小二乘法(PLSR)可以实现多对多线性回归建模,适合在自变量和因变量个数较多,数据间存在多重相关性,以及样本数量较少的情况下使用。一般情况下,偏最小二乘法与主成分分析法一样在建模过程中不需要选用存在的 n 个成分 $t_1, t_2, \cdots, t_n$,通过交叉有效性检验提取其中的 m 个主要成分($m \leqslant n$),从而实现性能较好预测模型的建立。

交叉有效性检验通过每次舍去第 i 个观测值($i = 1, 2, 3, \cdots, n$),用余下的 $n - 1$ 个观测值按最小二乘法进行建模,抽取 m 个成分后建立回归方程,然后把舍去的第 i 个观测点带入所建立的回归方程,得到 $y_j(j = 1, 2, \cdots, p)$ 在第 i 个观测点上的观测值 $\hat{y}_{ij}(m)$,对 $i = 1, 2, 3, \cdots, n$ 重复以上的验证,即得到抽取 m 个成分时第 j 个因变量 $y_j(j = 1, 2, \cdots, p)$ 的残差平方和[18]。

$$PRESS_j(m) = \sum_{i=1}^{n} [y_{ij} - \hat{y}_{ij}(m)]^2 (j = 1, 2, \cdots, p) \quad (1)$$

$Y = (y_1, y_2, \cdots, y_p)^T$ 的残差平方和为

$$PRESS(m) = \sum_{j=1}^{p} PRESS_j(m) \quad (2)$$

当 $PRESS(m)$ 达到最小时,对应的 m 即为所求的主成分变量个数,而本研究所涉及的输出变量只有一个,所以令上式 $p = 1$,即为单因变量 PLSR。

1.5　LSTM 神经网络模型

LSTM 神经网络模型是一种特殊的 RNN 神经网络模型,通过对长距离时间序列数据的学习,提高了模型预测的精度。LSTM 神经网络模型通过 sigmoid 函数和一个点乘运算组成一个具有信息选择功能的门,其中 sigmoid 函数的输出值在 0 ~ 1 之间,0 代表过滤信息,1 代表保留信息,LSTM 神经网络模型通过门能实现了信息的遗忘和记忆。LSTM 神经网络模型由很多细胞单元组成,每个细胞单元包括遗忘门、输入门和输出门,这些细胞单元使 LSTM 神经网络模型具有长期记忆信息的能力[19],LSTM 神经网络模型细胞单元结构如图 1 所示,其计算方法可表示为:

$$i_t = \sigma(w_i \times [h_{t-1}, x_t] + b_i) \quad (3)$$

$$f_t = \sigma(w_f \times [h_{t-1}, x_t] + b_f) \quad (4)$$

$$\tilde{c}_t = \tanh(w_c \times [h_{t-1}, x_t] + b_c) \tag{5}$$

$$c_t = f_t \times c_{t-1} + i_t \times \tilde{c}_t \tag{6}$$

$$o_t = \sigma(w_o \times [h_{t-1}, x_t] + b_o) \tag{7}$$

$$h_t = o_t \times \tanh(c_t) \tag{8}$$

式中：x_t 为本细胞单元的输入值；h_{t-1} 为上一细胞单元的输出值；i_t 为输入门；f_t 为遗忘门；o_t 为输出门；c_t 为 LSTM 神经网络模型细胞结构状态；w 为权重系数矩阵；b 为偏置项；σ 代表 sigmoid 函数；tanh 代表双曲正切激活函数[20]。该算法的具体计算步骤为：

（1）根据上一细胞单元的细胞状态 c_{t-1}，输出值 h_{t-1}，以及本细胞的输入值 x_t，按照公式（3）～（8）计算出本细胞单元的输出。

（2）按时间反向传播法，利用损失函数计算每个细胞单元的反向误差大小。

（3）根据反向误差，计算出对应的权重梯度。

（4）利用基于梯度的优化算法更新权重降低损失函数误差，直到 LSTM 神经网络模型收敛。

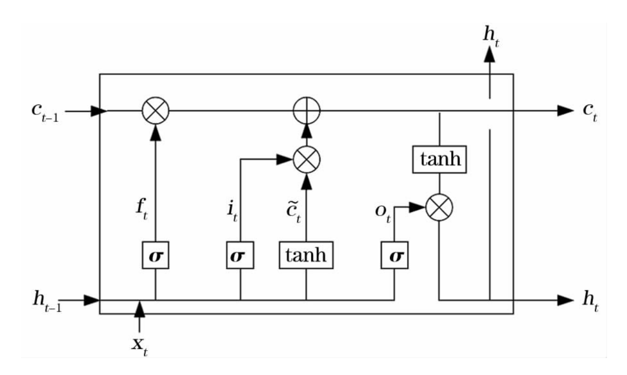


图 1　LSTM 神经网络模型细胞单元结构图

Fig. 1　Cell structure of LSTM neural network model

1.6　PLSR-LSTM 神经网络模型

首先根据 PLSR 交叉有效性原则从输入的 t 个自变量 $x_1, x_2, \cdots, x_t$ 中提取 m 个主成分 $t_1, t_2, \cdots, t_m$，经"归一化"处理后输入到 LSTM 神经网络模型的隐藏层，当数据 $t_1, t_2, \cdots, t_m$ 输入结束后，LSTM 神经网络模型的细胞单元对输入数据进行计算，并将隐藏层最终状态 h_m 传输到输出层，经"去归一"化处理后得到最终的预测结果 y，如图 2 所示。

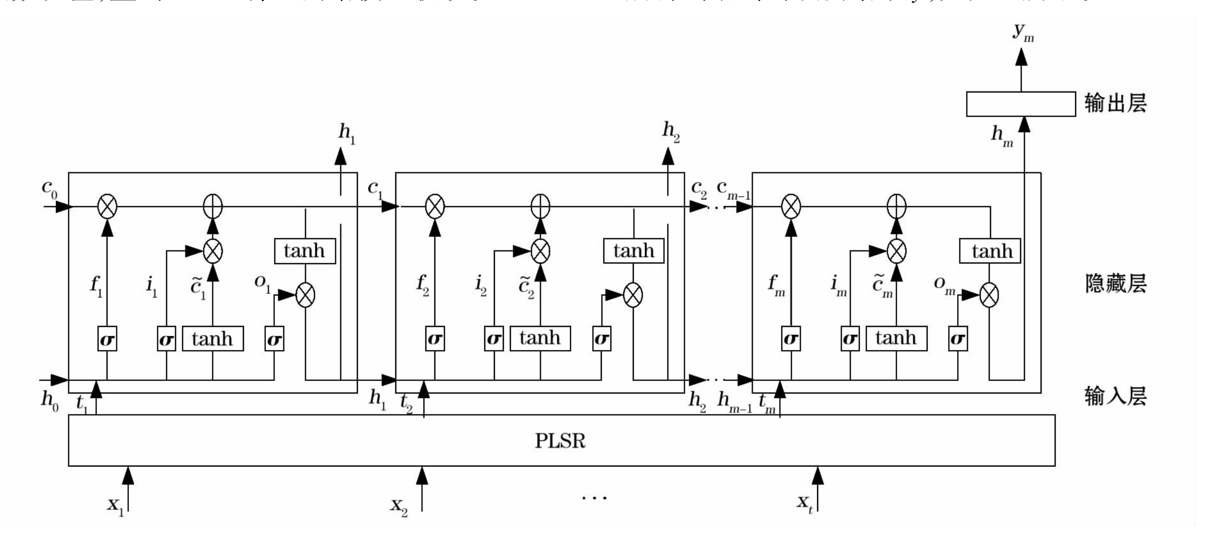


图 2　PLSR-LSTM 神经网络模型结构

Fig. 2　Structure of PLSR-LSTM neural network model

2　结果与分析

2.1　PLSR 主成分分析

本文将径流场所在流域的月降雨量 x_1、月水蚀性降雨量 x_2、月降雨强度 x_3、月径流量 x_4、径流场面积 x_5、径流场坡度 x_6、径流场土壤砂粒含量 x_7、径流场土壤粘粒含量 x_8 共 8 个因子作为自变量，月水蚀量 y 作为因变量，并对所用数据进行归一化处理。根据 PLSR 的交叉有效性原则对输入变量进行主成分分析以此降低由于影响因子共存而产生的相互重复信息，减少输入自变量个数并滤去噪音。

通过径流场输入变量 PLSR 主成分分析可以看出当主成分变量个数为 4 时，各径流场的残差平方和 *PRESS* 值达到最小，如图 3 所示。根据交叉有效性原则，本研究从选取的 8 个自变量中提取 4 个作为主成分输入 LSTM 神经网络模型。

本研究以月水蚀性降雨量 x_1、月降雨强度 x_2、月径流量 x_3、径流场坡度 x_4 为自变量，月水蚀量 y 为因变量，建立基于 PLSR-LSTM神经网络的人工草

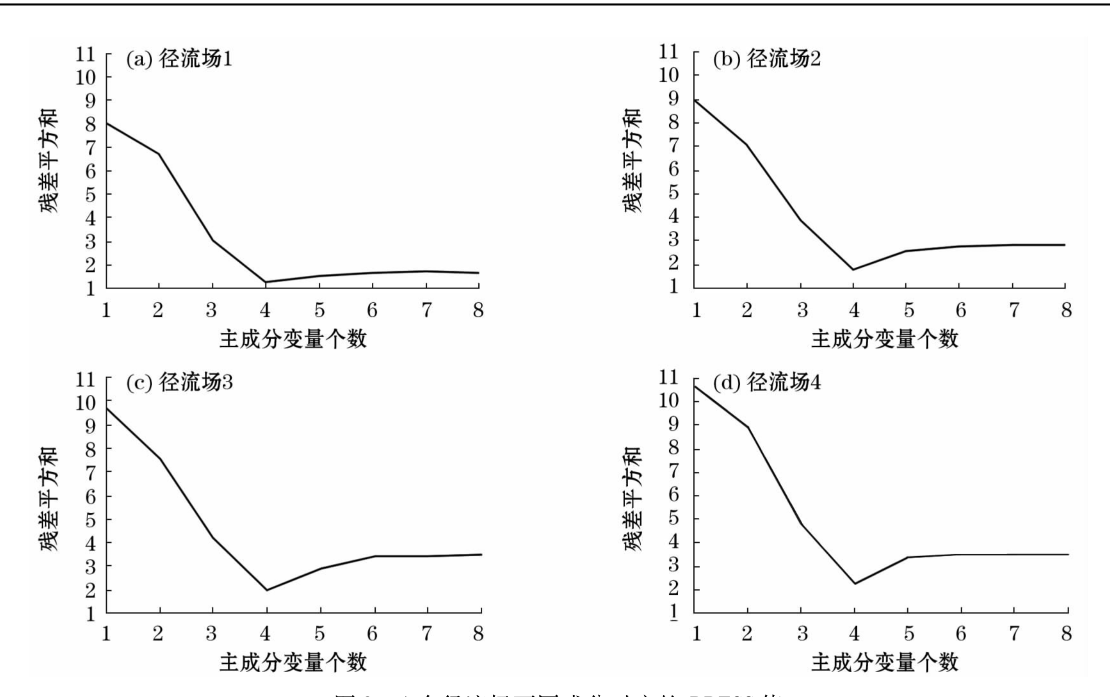


图 3　4 个径流场不同成分对应的 *PRESS* 值

Fig. 3　Variation of *PRESS* against number of components in four runoff fields

地土壤水蚀预测模型。

2.2　PLSR-LSTM 神经网络模型的训练及参数调整

基于 PLSR-LSTM 神经网络的土壤水蚀预测模型将定西市安家沟 2005—2016 年 1～12 月径流场试验数据按照训练样本占 70%,测试样本占 30%进行划分,其中 2005—2013 年 1～12 月试验数据作为训练样本对 PLSR-LSTM 神经网络模型进行训练,2014—2016 年 1～12 月试验数据作为测试样本用于 PLSR-LSTM 神经网络模型的有效性检验。

利用训练样本对 PLSR-LSTM 神经网络模型进行训练,采用正交化方法对权重系数矩阵进行初始化,将月水蚀性降雨量 x_1、月降雨强度 x_2、月径流量 x_3、径流场坡度 x_4 输入 PLSR-LSTM 神经网络模型得到月土壤水蚀量的预测值 y',并将该值与月土壤水蚀量实测值 y 的均方误差作为 PLSR-LSTM 神经网络模型的损失函数,损失函数如下所示:

$$loss = \frac{1}{N}\sum_{i=1}^{N}(y_i - y_i')^2 \qquad (9)$$

以损失函数 *loss* 最小为优化目标,在设定网络初始学习率 η、批尺寸以及最大迭代次数 step 的条件下,利用 Adam 下降优化算法经反复迭代更新网络权重系数矩阵,降低损失函数 *loss* 的值,直到 PLSR-LSTM 神经网络模型收敛,PLSR-LSTM 神经网络模型隐藏层参数如表 1 所示。

表 1　隐藏层网络参数

Tab. 1　Network parameters of hidden layer

参数名称	参数值
隐藏层数	3
最大迭代次数	600
批尺寸	72
损失函数	均方误差
下降方式	Adam 优化算法
学习速率	0.001
权值初始化方式	正交化初始化

2.3　PLSR-LSTM 神经网络模型的有效性与复杂度分析

为了验证 PLSR-LSTM 神经网络模型对陇中黄土丘陵沟壑区人工草地土壤水蚀量预测的有效性,本研究采用均方根误差(*RMSE*)、决定系数(R^2)和平均绝对误差(*MAE*)3 种统计指标来进行分析,其中 *RMSE*、*MAE* 越小,R^2 越接近 1,预测值与实测值间的误差越小,模型的拟合效果越好[21]。同时为了分析 PLSR-LSTM 神经网络模型在人工草地土壤水蚀量预测方面所具有的优势,本研究选用 BP[22]、RNN 以及 LSTM 等常见神经网络模型与该模型的预测结果进行对比,结果如表 2 所示。

PLSR-LSTM 神经网络模型对 4 个径流场土壤水蚀量预测结果的 *RMSE* 分别为 0.234、0.294、0.321、0.368;*MAE* 分别为 0.246、0.318、0.348、0.381;R^2 分别为 0.985、0.982、0.977、0.965。3 种

表 2　几种常见神经网络预测模型有效性对比

Tab. 2　Comparison of effectiveness of several common neural network prediction models

径流场	模型	*RMSE*	*MAE*	R^2
径流场 1	PLSR-LSTM	0.234	0.246	0.985
	LSTM	0.354	0.371	0.973
	RNN	0.397	0.411	0.958
	BP	0.424	0.457	0.943
径流场 2	PLSR-LSTM	0.294	0.318	0.982
	LSTM	0.389	0.392	0.961
	RNN	0.418	0.439	0.944
	BP	0.465	0.493	0.935
径流场 3	PLSR-LSTM	0.321	0.348	0.977
	LSTM	0.416	0.435	0.956
	RNN	0.442	0.481	0.941
	BP	0.493	0.524	0.929
径流场 4	PLSR-LSTM	0.368	0.381	0.965
	LSTM	0.447	0.487	0.940
	RNN	0.562	0.590	0.921
	BP	0.586	0.609	0.906

指标值均优于其他 3 种神经网络模型的预测结果。研究结果表明，PLSR-LSTM 神经网络模型能够有效预测陇中黄土丘陵沟壑区人工草地的土壤水蚀量。

为了对 PLSR-LSTM 神经网络人工草地土壤水蚀预测模型的复杂度进行检验，本研究将 4 种神经网络模型的损失误差进行了对比，如图 4 所示。研究结果表明，4 种模型的损失误差变化趋势相似，但 PLSR-LSTM 神经网络模型的损失误差小于其他 3 种模型，且该模型的迭代次数达到 54 次时模型的损失误差趋于稳定，收敛速度高于其他 3 种神经网络模型。

2.4　PLSR-LSTM 神经网络模型土壤水蚀量预测分析

PLSR-LSTM 神经网络模型利用 2014—2016 年 1～12 月的月水蚀性降雨量、月降雨强度、月径流量、径流场坡度，预测当月的月土壤水蚀量，2014—2016 年月平均土壤水蚀量如图 5 所示。从月均土壤水蚀量来看，陇中黄土丘陵沟壑区人工草地土壤水蚀主要集中在 5～9 月，其中 8 月的水蚀量最大，其次是 7 月和 9 月，4 个坡度径流场的月水蚀量大小依次为：径流场 4 > 径流场 3 > 径流场 2 > 径流场 1。

将月土壤水蚀量求和转换为年土壤水蚀量，并计算年土壤水蚀量预测值与实测值的相对误差，如表 3 所示。PLSR-LSTM 神经网络模型预测的年土壤水蚀量接近于实测值，平均相对误差小于 4%。由此可见，将 PLSR 与 LSTM 神经网络模型结合可以实现陇中黄土丘陵沟壑区人工草地土壤水蚀量的准确预测。为了更加直观显示模型的预测结果，将 2014—2016 年土壤水蚀量预测结果与真实值进行对比如图 6 所示。从年土壤水蚀量大小来看，径流

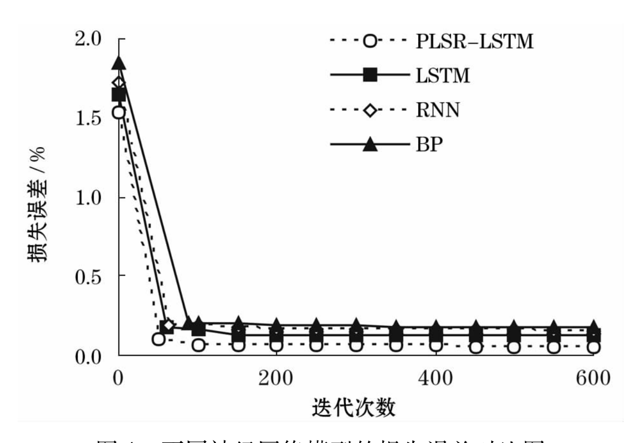


图 4　不同神经网络模型的损失误差对比图

Fig. 4　Comparison charts of loss error of different neural networks

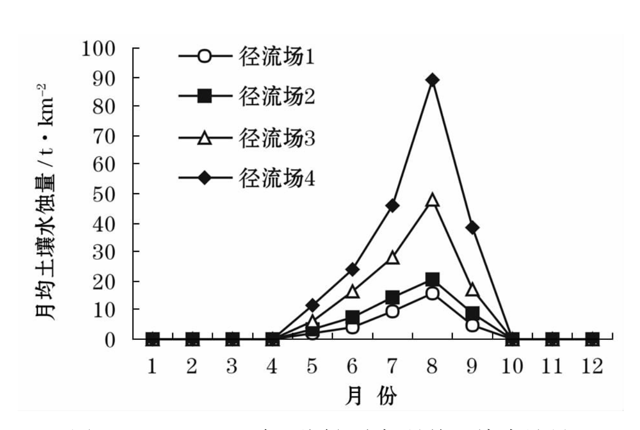


图 5　2014—2016 年不同径流场月均土壤水蚀量

Fig. 5　Monthly average soil water erosion in different runoff fields from 2014 to 2016

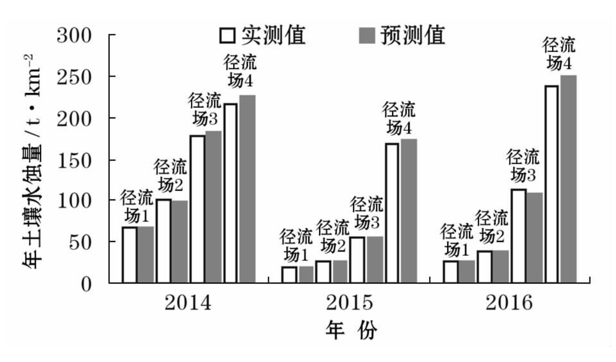


图 6　2014—2016 年土壤水蚀量实测值与预测值对比图

Fig. 6　Comparion chart of soil water erosion measured and predicted values from 2014 to 2016

表 3　2014—2016 年土壤水蚀量的实测值与预测值对比

Tab. 3　Comparison of soil erosion measured and predicted values from 2014 to 2016

径流场	2014 年			2015 年			2016 年		
	实测水蚀量	预测水蚀量	相对误差	实测水蚀量	预测水蚀量	相对误差	实测水蚀量	预测水蚀量	相对误差
径流场 1	67.8	66.4	2.06	18.6	18.8	1.08	26.2	25.8	1.53
径流场 2	100.8	98.6	2.18	27.6	27.2	1.45	38.8	37.6	3.09
径流场 3	176.8	182.8	3.39	56.2	55.2	1.78	112.6	109.2	3.02
径流场 4	215.8	225.6	4.54	167.0	172.2	3.13	238.4	249.8	4.78

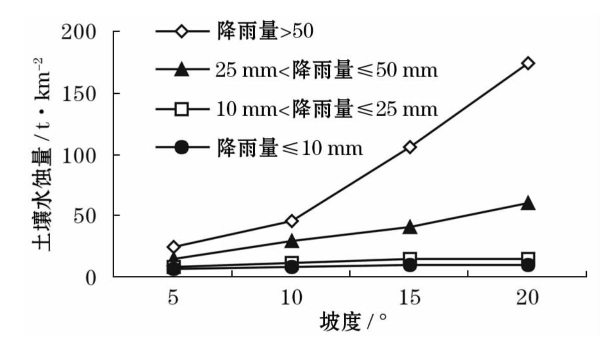


图 7　不同降雨量条件下人工草地水蚀量与坡度间的关系

Fig. 7　Relationship between water erosion and slope in artificial grassland with different rainfall

场 4 > 径流场 3 > 径流场 2 > 径流场 1,研究结果表明,坡度对土壤水蚀的影响较为明显。

不同降雨量条件下土壤水蚀量与坡度的关系,如图 7 所示。不同坡度条件下,降雨量≤10 mm 时引起的土壤水蚀量最小,且坡度增加对土壤水蚀量的影响不大;当 10 mm < 降雨量≤25 mm 时,5° ~ 20°坡度区间的土壤水蚀量随着坡度的增加呈小幅度增加,增幅并不明显;当 25 mm < 降雨量≤50 mm 时,随着坡度的增加,土壤水蚀量增幅加大;当降雨量 > 50 mm 时,土壤水蚀量随坡度增加大幅度增长,其中坡度大于 10°时,土壤水蚀量的增幅最为明显。

3　结 论

土壤水蚀过程是一种复杂的非线性过程,由于 LSTM 神经网络模型具有很好的非线性映射能力,因此该模型可以被应用于土壤水蚀量预测,但由于模型精度对学习速率取值敏感,使得学习速率如果设置的不合理可能造成模型预测结果出现欠拟合或过拟合现象,影响模型对土壤水蚀预测的准确度。将 PLSR 主成分提取与 LSTM 神经网络模型结合,充分发挥各自优点,通过主成分分析剔除观测数据的重复信息,在有效降低数据间相关性的同时,也解决了 LSTM 神经网络模型对实测数据要求过高的问题,简化了 LSTM 神经网络模型的系统结构,提高了模型的收敛速度和可靠性。PLSR-LSTM 神经网络模型土壤水蚀量预测结果的平均相对误差小于 4%,*RMSE*、*MAE* 和 R^2 等统计指标均明显优于 BP、RNN、LSTM 等常见神经网络模型。研究结果表明,该模型能较好的实现陇中黄土丘陵沟壑区人工草地土壤水蚀的准确预测。

降雨是导致陇中黄土丘陵沟壑区人工草地坡面产生土壤水蚀的主要因素之一,同时坡度对人工草地土壤水蚀的影响也比较明显,坡度的增加会导致土壤水蚀量的增加。研究结果表明,当降雨量小于 25 mm 时,人工草地土壤水蚀量并不会随坡度增加而明显增长,但当降雨超过 25 mm 时,人工草地土壤水蚀量会随坡度明显增加。

本研究所采用的输入变量均为容易获取或测定的数据,但在预测过程中没有考虑植被、地形等其他因素对土壤水蚀预测的影响。因此,本研究采用的数据并不能准确地还原土壤水蚀形成区的要素状况,使得土壤水蚀预测过程存在误差,在后期的研究过程中,可以通过增加新的约束条件来进一步提高模型的预测精度,为陇中黄土丘陵沟壑区土壤水蚀的准确预报提供有效的技术支持。

参考文献(References)

[1] 王权威,唐莉. 基于 ABC-BP 的土壤侵蚀量预报模型研究[J]. 水利发电,2017,43(9):1 – 4. [WANG Quanwei, TANG Li. Study on soil erosion prediction model based on artificial bee colony algorithm and BP neural network[J]. Water Power, 2017, 43(9):1 – 4.]

[2] 卫伟,陈利顶,傅伯杰,等. 半干旱黄土丘陵沟壑区降水特征值和下垫面因子影响下的水土流失规律[J]. 生态学报,2006,26(11):3847 – 3853. [WEI Wei, CHEN Liding, FU Bojie, et al. Mechanism of soil and water loss under rainfall and earth surface characteristics in a semiarid loess hilly area[J]. Acta Ecologica Sinica, 2006, 26(11):3847 – 3853.]

[3] 张兴奇,顾礼彬,张科利,等. 坡度对黔西北地区坡面产流产沙的影响[J]. 水土保持学报,2015,29(4):18-22,72. [ZHANG Xingqi,GU Libin,ZHANG Keli,et al. Impacts of slope gradient on runoff and sediment in northwest Guizhou[J]. Journal of Soil and Water Conservation,2015,29(4):18-22,72.]

[4] 和继军,蔡强国,刘松波. 次降雨条件下坡度对坡面产流产沙的影响[J]. 应用生态学报,2012,23(5):1263-1268. [HE Jijun,CAI Qiangguo,LIU Songbo. Effects of slope gradient on slope runoff and sediment yield under different single rainfall conditions[J]. Chinese Journal of Applied Ecology,2012,23(5):1263-1268.]

[5] 杨春霞,肖培青,甄斌,等. 野外不同被覆坡面产流产沙特征[J]. 水土保持学报,2012,26(4):28-36. [YANG Chunxia,XIAO Peiqing,ZHEN Bin,et al. Characteristics on runoff and sediment yield in field prototype and different coverage slope[J]. Journal of Soil and Water Conservation,2012,26(4):28-36.]

[6] 胡琳,苏静,桑永枝,等. 陕西省降雨侵蚀力时空分布特征[J]. 干旱区地理,2014,37(6):1101-1107. [HU Lin,SU Jing,SANG Yongzhi,et al. Spatial and temporal characteristics of rainfall erosivity in Shaanxi Province[J]. Arid Land Geography,2014,37(6):1101-1107.]

[7] 王凯,陈璐,马金辉,等. TRMM 数据在中国降雨侵蚀力计算中的应用[J]. 干旱区地理,2015,38(5):948-959. [WANG Kai,CHEN Lu,MA Jinhui,et al. Calculation of rainfall erosivity in China with TRMM data[J]. Arid Land Geography,2015,38(5):948-959.]

[8] RENARD K G,FOSTER G R,WEESIES G A,et al. RUSLE:Revised universal soil loss equation[J]. Journal of Soil and Water Conservation,1991,46(1):30-33.

[9] 王略,屈创,赵国栋. 基于中国土壤流失方程模型的区域土壤侵蚀定量评价[J]. 水土保持通报,2018,38(1):122-125. [WANG Lue,QU Chuang,ZHAO Guodong. Quantitative assessment of regional soil erosion based on Chinese soil loss equation model[J]. Bulletin of Soil and Water Conservation,2018,38(1):122-125.]

[10] LANGK V M,KARLSSON L,LOUTFI A,et al. A review of unsupervised feature learning and deep learning for time-series modeling[J]. Pattern Recognition Letters,2014,42(1):11-24.

[11] LECUN Y,BENGIO Y,HINTON G. Deep learning[J]. Nature,2015,521(7553):436-444.

[12] 范竣翔,李琦,朱亚杰,等. 基于 RNN 的空气污染时空预报模型研究[J]. 测绘科学,2017,42(7):76-83. [FAN Junxiang,LI Qi,ZHU Yajie,et al. A spatio-temporal prediction framework for air pollution based on deep RNN[J]. Science of Surveying and Mapping,2017,42(7):76-83.]

[13] 王鑫,吴际,刘超,等. 基于 LSTM 循环神经网络的故障时间序列预测[J]. 北京航空航天大学学报,2018,44(4):772-782. [WANG Xin,WU Ji,LIU Chao,et al. Exploring LSTM based recurrent neural network for failure time series prediction[J]. Journal of Beijing University of Aeronautics and Astronautics,2018,44(4):772-782.]

[14] 李世欣,温建,邵孝侯,等. 偏最小二乘法与人工神经网络耦合的小流域产沙模型[J]. 海河大学学报,2010,38(2):149-152. [LI Shixin,WEN Jian,SHAO Xiaohou,et al. Sediment yield model for small watersheds based on coupling of partial least square regression and artificial neural network[J]. Journal of Hohai University,2010,38(2):149-152.]

[15] 聂敏,刘志辉,刘洋,等. 基于 PCA 和 BP 神经网络的径流预测[J]. 中国沙漠,2016,36(4):1144-1151. [NIE Min,LIU Zhihui,LIU Yang,et al. Runoff forecast based on principal component analysis and BP neural network[J]. Journal of Desert Research,2016,36(4):1144-1151.]

[16] 张朋,卜崇峰,杨永胜,等. 基于 CCA 的坡面尺度生物结皮空间分布[J]. 生态学报,2015,35(16):5412-5420. [ZHANG Peng,BU Chongfeng,YANG Yongsheng,et al. The slope scale distribution regularity of biological soil crusts based on CCA[J]. Acta Ecologica Sinica,2015,35(16):5412-5420.]

[17] 谢云,刘宝元,章文波. 侵蚀性降雨标准研究[J]. 水土保持学报. 2000,12(4):6-11. [XIE Yun,LIU Baoyuan,ZHANG Wenbo. Study on standard of erosive rainfall[J]. Journal of Soil and Water Conservation,2000,12(4):6-11.]

[18] 张东平,孙克勤,潘效军,等. 烟气脱硫系统增压风机仿真模型开发及应用[J]. 环境工程学报,2011,5(9):2082-2086. [ZHANG Dongping,SUN Keqin,PAN Xiaojun,et al. Development and application of boost fan simulation model for flue gas desulfurization system[J]. Chinese Journal of Environmental Engineering,2011,5(9):2082-2086.]

[19] GREFF K,SRIVASTAVA R K,KOUTNIK J,et al. LSTM:A search space odyssey[J]. IEEE Transactions on Neural Networks & Learning Systems,2016,(99):1-11.

[20] 王旭东,严珂,陆慧娟,等. LSTM 的单变量短期家庭电力需求预测[J]. 中国计量大学学报,2018,29(2):142-147. [WANG Xudong,YAN Ke,LU Huijuan,et al. Short-term household electricity demand forecast based on LSTM single variable[J]. Journal of China University of Metrology,2018,29(2):142-147.]

[21] ABUDU S,CUI C L,KING J P,et al. Comparison of performance of statistical models in forecasting monthly streamflow of Kizil River China[J]. Water Science and Engineering,2010,3(3):269-281.

[22] 刘婷,邵景安. 基于 BP 神经网络的三峡库区土壤侵蚀强度模拟[J]. 自然资源学报,2018,33(4):669-683. [LIU Ting,SHAO Jing'an. Simulation of soil erosion intensity in the three gorges reservoir area using BP neural network[J]. Journal of Natural Resources,2018,33(4):669-683.]

Testing a soil water erosion predictive model in an artificial grassland in the hill-gully Loess Plateau

WANG Jun[1,2], LI Guang[1], NIE Zhi-gang[2], LIU Qiang[2]

(1 *College of Forestry, Gansu Agricultural University, Lanzhou* 730000, *Gansu, China*;

2 *College of Information Science and Technology, Gansu Agricultural University, Lanzhou* 730000, *Gansu, China*)

Abstract: Erosion caused by soil water is widespread in the hill-gully Loess Plateau of central Gansu, China, but it is difficult to observe and predict. The objectives of this study were to establish a predictive model for soil water erosion based on a partial least squares algorithm (PLSR) and long short-term memory (LSTM), and to evaluate the model's effectiveness using a back-propagation neural network, a recurrent neural network, and a LSTM neural network. The results were verified with experimental data from a runoff field of artificial grassland in the soil and water conservation experimental stations in Anjiagou, Dingxi City, from 2005 to 2016 (1 – 12 months). Input factors for the model were monthly precipitation, monthly erosive precipitation, monthly runoff, monthly rainfall intensity, runoff area, slope of runoff field, soil sand content, and clay sand content , from which the model predicted monthly soil water erosion. The results revealed that the original eight input factors could be reduced to four by the PLSR algorithm in the prediction model, alleviating the requirement of the LSTM neural network model for a large number of samples. Our method improved both the prediction precision and the convergence rate of the soil water erosion model in artificial grassland by combining a PLSR algorithm with a LSTM neural network. The average relative error of our prediction results was less than 4% , and the correlation coefficients were higher for our model than for the three neural network models used for comparison. In addition, iteration times, root mean square error, and mean absolute error were lower than for the other neural network models. We found that slope had a significant effect on soil water erosion, but only when the monthly rainfall was more than 25 mm. Our PLSR-LSTM neural network model could be used to accurately predict soil water erosion in artificial grassland in the hill-gully Loess Plateau of central Gansu, and it could provide a new path toward accurate prediction of soil erosion in this area.

Key words: hill-gully loess plateau; artificial grassland; long short-term memory (LSTM) recurrent neural network; soil water erosion prediction